



**HAL**  
open science

## The cosipy library: COSI's high-level analysis software

Israel Martinez-Castellanos, Savitri Gallego, Chien-You Huang, Chris Karwin, Carolyn Kierans, Jan Peter Lommler, Saurabh Mittal, Michela Negro, Eliza Neights, Sean Pike, et al.

### ► To cite this version:

Israel Martinez-Castellanos, Savitri Gallego, Chien-You Huang, Chris Karwin, Carolyn Kierans, et al.. The cosipy library: COSI's high-level analysis software. International Cosmic-Ray Conference, Jul 2023, Nagoya, Japan. 10.48550/arXiv.2308.11436 . hal-04300116

**HAL Id: hal-04300116**

**<https://hal.science/hal-04300116>**

Submitted on 22 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

## The cosipy library: COSI's high-level analysis software

Israel Martinez-Castellanos,<sup>u,f,\*</sup> Savitri Gallego,<sup>r</sup> Chien-You Huang,<sup>q</sup> Chris Karwin,<sup>f</sup> Carolyn Kierans,<sup>f</sup> Jan Peter Lommler,<sup>r</sup> Saurabh Mittal,<sup>h</sup> Michela Negro,<sup>d</sup> Eliza Neights,<sup>f,l</sup> Sean N. Pike,<sup>b</sup> Yong Sheng,<sup>c</sup> Thomas Siegert,<sup>h</sup> Hiroki Yoneda,<sup>h</sup> Andreas Zoglauer,<sup>a</sup> John A. Tomsick,<sup>a</sup> Steven E. Boggs,<sup>b,a</sup> Dieter Hartmann,<sup>c</sup> Marco Ajello,<sup>c</sup> Eric Burns,<sup>d</sup> Chris Fryer,<sup>e</sup> Alexander Lowell,<sup>a</sup> Julien Malzac,<sup>g</sup> Jarred Roberts,<sup>b</sup> Pascal Saint-Hilaire,<sup>a</sup> Albert Shih,<sup>f</sup> Clio Sleator,<sup>i</sup> Tadayuki Takahashi,<sup>j</sup> Fabrizio Tavecchio,<sup>k</sup> Eric Wulf,<sup>i</sup> Jacqueline Beechert,<sup>a</sup> Hannah Gulick,<sup>a</sup> Alyson Joens,<sup>a</sup> Hadar Lazar,<sup>a</sup> Juan Carlos Martinez Oliveros,<sup>a</sup> Shigeki Matsumoto,<sup>j</sup> Tom Melia,<sup>j</sup> Mark Amman,<sup>m</sup> Dhruv Bal,<sup>c</sup> Peter von Ballmoos,<sup>g</sup> Hugh Bates,<sup>c</sup> Markus Böttcher,<sup>n</sup> Andrea Bulgarelli,<sup>o</sup> Elisabetta Cavazzuti,<sup>p</sup> Hsiang-Kuang Chang,<sup>q</sup> Claire Chen,<sup>a</sup> Che-Yen Chu,<sup>q</sup> Alex Ciabattini,<sup>o</sup> Luigi Costamante,<sup>p</sup> Lente Dreyer,<sup>n</sup> Valentina Fioretti,<sup>o</sup> Francesco Fenu,<sup>p</sup> Giancarlo Ghirlanda,<sup>k</sup> Eric Grove,<sup>i</sup> Pierre Jean,<sup>g</sup> Nikita Khatiya,<sup>c</sup> Jürgen Knödlseder,<sup>g</sup> Martin Krause,<sup>s</sup> Mark Leising,<sup>c</sup> Tiffany R. Lewis,<sup>f,w</sup> Lea Marcotulli,<sup>t</sup> Samer Al Nussirat,<sup>a</sup> Kazuhiro Nakazawa,<sup>v</sup> Uwe Oberlack,<sup>r</sup> David Palmore,<sup>d</sup> Gabriele Panebianco,<sup>o</sup> Nicolo Parmiggiani,<sup>o</sup> Tyler Parsotan,<sup>f</sup> Field Rogers,<sup>a</sup> Hester Schutte,<sup>n</sup> Alan P. Smale,<sup>f</sup> Jacob Smith,<sup>i</sup> Aaron Trigg,<sup>d</sup> Tonia Venters,<sup>f</sup> Yu Watanabe<sup>j</sup> and Haocheng Zhang<sup>f</sup>

<sup>a</sup>Space Sciences Laboratory, 7 Gauss Way, University of California, Berkeley CA 94720-7450, USA

<sup>b</sup>Department of Astronomy & Astrophysics, UC San Diego, 9500 Gilman Drive, La Jolla CA 92093, USA

<sup>c</sup>Clemson University, South Carolina, USA

<sup>d</sup>Louisiana State University, Baton Rouge, LA 70803, USA

<sup>e</sup>Los Alamos National Laboratory, New Mexico, USA

<sup>f</sup>NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA

<sup>g</sup>Institut de Recherche en Astrophysique et Planetologie, 9, avenue du Colonel Roche BP 44346 31028 Toulouse Cedex 4, France

<sup>h</sup>Institut für Theoretische Physik und Astrophysik, Universität Würzburg, Campus Hubland Nord, Emil-Fischer-Str. 31, 97074, Würzburg, Germany

<sup>i</sup>U.S. Naval Research Laboratory, 4555 Overlook Ave., SW Washington, DC 20375, USA

<sup>j</sup>Kavli Institute for the Physics and Mathematics of the Universe, The University of Tokyo, Japan

<sup>k</sup>Istituto Nazionale di Astrofisica, Merate, Italy

<sup>l</sup>George Washington University, USA

<sup>m</sup>Independent, USA

\*Speaker

<sup>n</sup>Centre for Space Research, North-West University, Potchefstroom 2520, South Africa

<sup>o</sup>Istituto Nazionale di Astrofisica, Bologna, Italy

<sup>p</sup>Italian Space Agency, Italy

<sup>q</sup>Institute of Astronomy, National Tsing Hua University, Guangfu Rd., Hsinchu City, 300044, Taiwan

<sup>r</sup>Institut für Physik & Exzellenzcluster PRISMA<sup>+</sup>, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany

<sup>s</sup>Centre for Astrophysics Research, Department of Physics, Astronomy and Mathematics, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK

<sup>t</sup>Yale University, USA

<sup>u</sup>University of Maryland, USA

<sup>v</sup>Nagoya University, Japan

<sup>w</sup>NASA Postdoctoral Program Fellow

E-mail: [israel.martinez@nasa.gov](mailto:israel.martinez@nasa.gov)

The Compton Spectrometer and Imager (COSI) is a selected Small Explorer (SMEX) mission launching in 2027. It consists of a large field-of-view Compton telescope that will probe with increased sensitivity the under-explored MeV gamma-ray sky (0.2-5 MeV). We will present the current status of cosipy, a Python library that will perform spectral and polarization fits, image deconvolution, and all high-level analysis tasks required by COSI's broad science goals: uncovering the origin of the Galactic positrons, mapping the sites of Galactic nucleosynthesis, improving our models of the jet and emission mechanism of gamma-ray bursts (GRBs) and active galactic nuclei (AGNs), and detecting and localizing gravitational wave and neutrino sources. The cosipy library builds on the experience gained during the COSI balloon campaigns and will bring the analysis of data in the Compton regime to a modern open-source likelihood-based code, capable of performing coherent joint fits with other instruments using the Multi-Mission Maximum Likelihood framework (3ML). In this contribution, we will also discuss our plans to receive feedback from the community by having yearly software releases accompanied by publicly-available data challenges.

## 1. Overview

The Compton Spectrometer and Imager (COSI)[1] is a compact Compton telescope sensitive to gamma rays between 200 keV and 5 MeV. COSI is a survey instrument with a large field of view  $\rightarrow \pi$  sr— excellent energy resolution  $\ll 1.2\%$  ( $\ll 0.07\%$ ) at 511 keV (1.157 MeV)— capable of imaging with a good angular resolution  $\ll 4.1^\circ$  ( $\ll 2.1^\circ$ ) at 511 keV (1.809 MeV)— and performing polarization measurements. COSI will reduce the so-called “MeV gap”, the low-energy part of the gamma-ray spectrum that despite its great scientific potential is currently under explored. COSI will uncover the origin of the Galactic positrons, advance our understanding of how the elements were formed, and probe the physics of extreme environments and multi-messenger sources.

The COSI team is currently developing *cosipy*, a Python library that will be capable of performing all the high-level task expected for COSI: imaging, spectral analysis and polarimetry. On this contribution we first introduce, in Section 2, the basic principles of Compton data analysis. This is then followed by the design of *cosipy* and its current status. We also include our plans for the gradual improvement of *cosipy* and public data challenges, leading up to the expected launch of COSI in 2027.

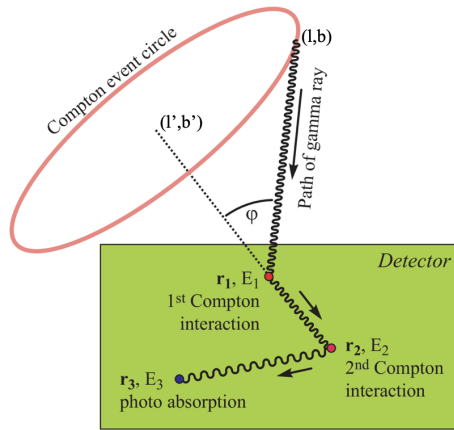
## 2. Compton data analysis

In the MeV regime photons interact with the detector predominantly through Compton scattering. As shown in Fig. 1, the incoming photon results in a series of energy deposits in the detector until it is finally absorbed due to the photoelectric effect. Compact telescope like COSI cannot resolve the order of the interactions based on their timing, and therefore it must be derived based on the kinematics of Compton scattering. Once this is determined, the scattering angle  $\phi$  and direction ( $l', b'$ ) of the outgoing photon in the first interaction define a circle in the sky that contains the direction of the incoming photon ( $l, b$ ). The set of parameters  $\{\phi, l', b'\}$  is usually called the Compton data space (CDS). Although the following interactions also contain useful information, the CDS is the minimum information needed to perform imaging and polarimetry.

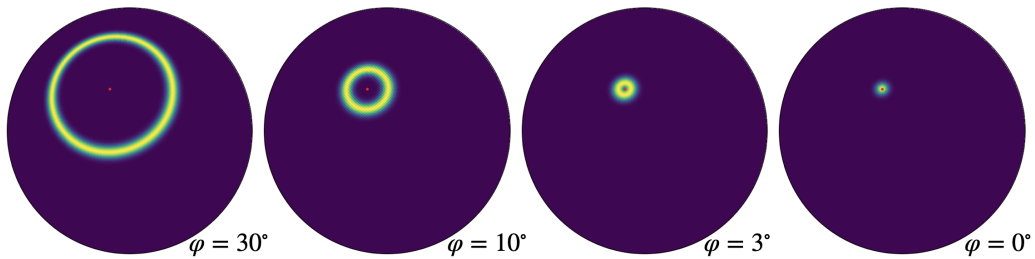
Since the radius of the Compton circle is fully constrained by the scattering angle  $\phi$ , the combination over all possible values of  $\phi$  define a cone in the three-dimensional CDS. The walls of this cones are broaden due to instrumental effects, such as the unknown kinetic energy of the electron that participated in the first interaction, the finite energy resolution, missed interactions due to the detection threshold, etc. The distribution of the broaden Compton cone is the point spread function (PSF) of the instrument, as shown in Fig. 2.

A source can be localized by combining the information from multiple events, as shown in Fig 3. The imaging of the sky involves multiple sources and backgrounds events, not shown in Fig. 3a, that require the use of complex deconvolution algorithms. All these calculations are performed in the CDS where the Compton cones for each source can be distinguished, as exemplified in Fig. 3b.

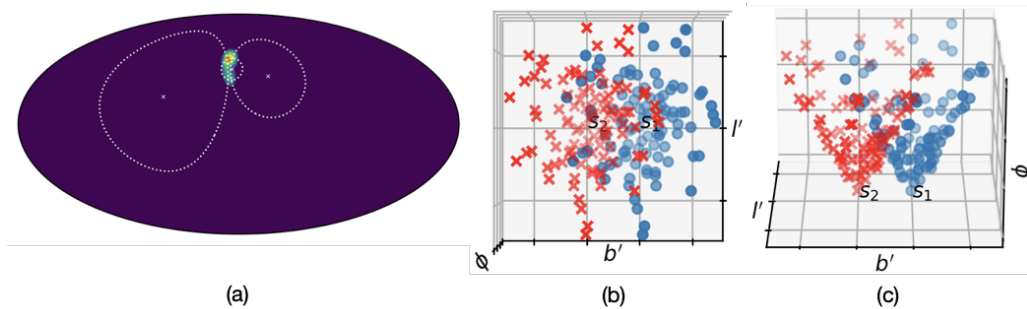
Polarimetry is possible thanks to the azimuthal dependence of Compton scattering in response to a linearly polarized source. As shown in Fig. 4, the distribution of events in the Compton cone is modulated by the polarization angle—determining the phase— and the polarization fraction—proportional to the modulation factor.



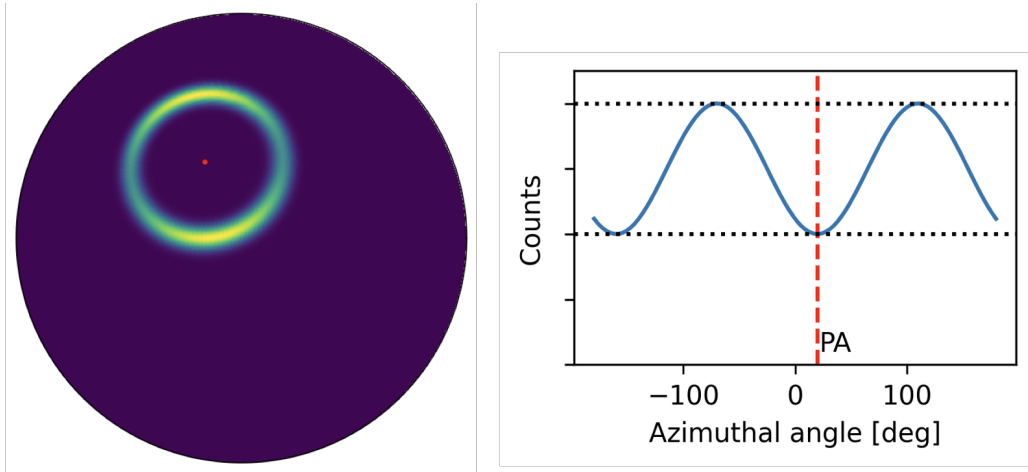
**Figure 1:** COSI measures the location and the energy deposited by the different interactions of a gamma ray with the detector. During reconstruction, this path is deduced based on the kinematics of Compton scattering, constraining the direction of the incoming photon to a circle in the sky, centered in the direction of the first scattered gamma ray  $(l', b')$  and with a radius equal to the scattering angle of the first interaction  $\phi$ .



**Figure 2:** Different slices of a simplified point spread function (PSF) in the Compton data space (CDS). The PSF was sliced along the  $\phi$ -axis (scattering angle) and shown in an orthographic projection of the sphere describing the direction of the outgoing scattered photon. The red dot shows the true location of the incoming photon.



**Figure 3:** (a) A simplified example demonstrating the localization of a source (red dot) based on three detected events (white crosses, with Compton circles shown with dashed lines). (b,c) Events produced by two sources ( $s_1$  and  $s_2$ ) distributed in the Compton Data Space (CDS). Views from the top (b) and side (c). By keeping track of the scattered angle  $\phi$  the cones can be discriminated and the image deconvolved.



**Figure 4:** Left: simplified point spread function (PSF) for a linearly polarized sources. Right: projection onto the azimuth angle. The ratio between the black dotted lines is proportional to the polarization fraction. The phase is determined by the polarization angle (PA).

### 3. The cosipy library

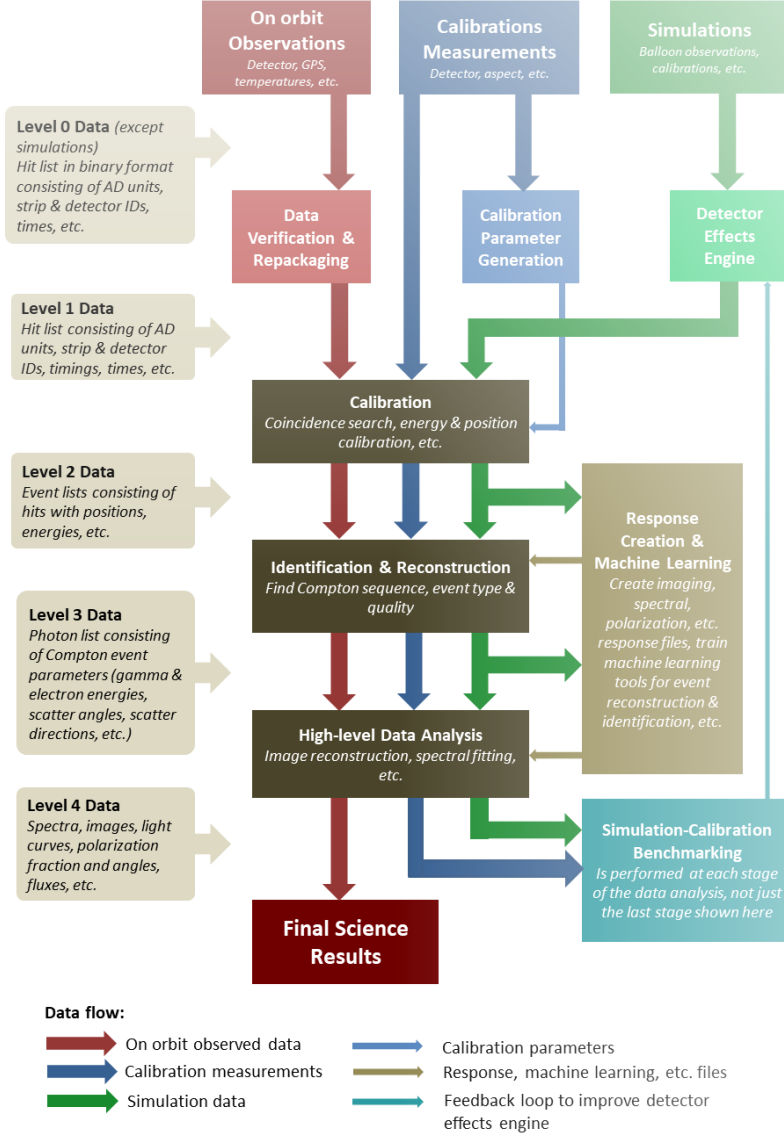
Although the basic principles of Compton data analysis can be applied to approximate the spectrum, location and polarization of source, real data presents significantly more challenges. The large number of events, high and variable backgrounds, instrument artifacts, and the continuous rotation of the spacecraft, all call for a more complex and robust high-analysis software. The time-dependent spectra and polarization, and correlations between energy, polarization and direction, also complicate the analysis.

The cosipy library is a component of a series of software packages called *cositools*, and the last step of the COSI calibration and analysis pipeline, shown in 5. It builds upon the development during the COSI balloon campaigns. For a full description of the calibration and pipelines see [2, 3].

#### 3.1 Current design and status

The cosipy library follows a likelihood-based approach. It uses a forward folding technique to compare the expected measurements to the data —energy  $E_m$ , scattering angle  $\phi$ , and outgoing photon direction  $(l', b')$ , which together we will call the data space  $\Omega$ — given a sky model hypothesis —i.e. a distribution of sources, each with a given spectrum. The relationship between the expected measurements and physical photon parameters — incoming direction  $(l, b)$ , energy  $E_i$  and polarization angle  $PA$ , which we will call the photon parameter space  $\Lambda$ — is handled by the detector response, and encoded in a multi-dimensional matrix.

The likelihood can be written as:

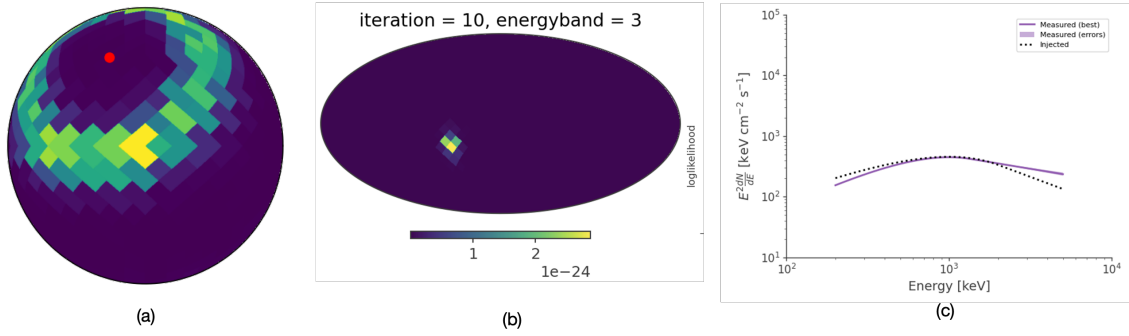


**Figure 5:** The cosipy library will perform the *High-level Data Analysis* tasks. It is the last step of a series of software packages that compose the COSI calibration and analysis pipeline.

$$\log \mathcal{L}(\mathbf{s}, \mathbf{b}) = \sum_{\Delta\Omega_i} \log P(n_j | \lambda_j) \quad (1)$$

$$\lambda_i(\mathbf{s}, \mathbf{b}) = \int_{\Delta\Omega_i} \left( \int A_{\text{eff}}(\Lambda) P(\Omega | \Lambda) F(\mathbf{s}, \Lambda) d\Lambda + \text{Bkg}(\mathbf{b}) \right) d\Omega,$$

where  $\mathbf{s}$  and  $\mathbf{b}$  are the signal and background parameters to fit;  $P(n|\lambda)$  is the Poisson probability distribution; the term  $A_{\text{eff}}(\Lambda) P(\Omega|\Lambda)$  is the detector response, equal to the effective area times the probability of a detected event with physical parameters  $\Lambda$  to have the measured parameters  $\Omega$ ; and  $F$  is the flux for a given location in the sky at a given energy and polarization.



**Figure 6:** Results from an early version of cosipy. (a) Simulated binned data, using a coarse grid, showing the location of a source (red dot) and the Compton circle. (b) Deconvolved image of the source using the Richardson-Lucy algorithm. (c) Fitted spectrum of the source using the 3ML framework.

The detector response is obtained from event by event Monte Carlo simulations using MEGALib [4]. All passive and active components are carefully modeled, and the detector effects simulated, in order to match benchmark calibration measurements.

The cosipy library is compatible with the Multi-Mission Maximum Likelihood framework (3ML) [5] which provides a common interface between multiple instruments. This is achieved by the use of plugins that receive a source model in a common format, and output the corresponding likelihood. This allows us to gain access to the maximum likelihood and Bayesian algorithms provided by 3ML in order to do spectral and polarization analysis.

Imaging deconvolution can also be performed by maximizing the likelihood in Eq. 1 over multiple hypothetical sources in a grid fully covering the sky. Due to the high dimensionality of the problem we use the Richardson-Lucy algorithm [6, 7] which consists of an interactive method that converges to the maximum likelihood solution.

We have preliminary versions of the detectors response handling, likelihood calculation, imaging and spectral analysis. Selected results are shown in Fig. 6. During these early stages of development we are using a simplified detector response and coarse binning in order to test and profile the code. Currently we are working on optimizing the CPU and memory usage such that the computation can scale well to requirements of the flight version.

### 3.2 Data challenges and future plans

We will progressively add new capabilities and improvements to cosipy, leading up to the launch of COSI in 2027. Starting in 2024, we will organize public data challenges with a yearly cadence, consisting of simulated data sets for sources and science topics of interest. This will build upon the COSI balloon data challenge released in 2023 [8]. We want to obtain feedback from the community to find issues and identify missing features required to satisfy all COSI science goals as well as to expand its portfolio.

Some of the plans for the upcoming releases include:

- Simultaneous location, spectral and polarization fits.
- Time-dependent polarization analysis.



- Implementation of alternative imaging deconvolution algorithms, such as Multiresolution Regularized Expectation Maximization, and machine-learning-based image creation.

#### 4. Conclusion

The cosipy library will perform all high-level analysis tasks to satisfy the scientific goals of the COSI mission, an MeV gamma-ray telescope launching in 2027. Currently under development, cosipy follows a modern likelihood-based approach, is open source, and compatible with the Multi-Mission Maximum Likelihood framework (3ML). Starting in 2024, we will make yearly releases accompanied by a public data challenge in order to obtain feedback from the community.

#### References

- [1] J. Tomsick, A. Zoglauer, C. Sleator, H. Lazar, J. Beechert, S. Boggs et al., *The Compton Spectrometer and Imager*, in *Bulletin of the American Astronomical Society*, vol. 51, p. 98, Sept., 2019 [[1908.04334](#)].
- [2] J. Beechert, H. Lazar, S.E. Boggs, T.J. Brandt, Y.-C. Chang, C.-Y. Chu et al., *Calibrations of the Compton Spectrometer and Imager*, *Nuclear Instruments and Methods in Physics Research A* **1031** (2022) 166510 [[2203.00695](#)].
- [3] A. Zoglauer, T. Siegert, A. Lowell, B. Mochizuki, C. Kierans, C. Sleator et al., *COSI: From Calibrations and Observations to All-sky Images*, *arXiv e-prints* (2021) [arXiv:2102.13158](#) [[2102.13158](#)].
- [4] A. Zoglauer, R. Andritschke and F. Schopper, *MEGALib The Medium Energy Gamma-ray Astronomy Library*, *New Astronomy Reviews* **50** (2006) 629.
- [5] G. Vianello, R.J. Lauer, P. Younk, L. Tibaldo, J.M. Burgess, H. Ayala et al., *The multi-mission maximum likelihood framework (3ml)*, *arXiv preprint arXiv:1507.08343* (2015) .
- [6] L.B. Lucy, *An iterative technique for the rectification of observed distributions*, *Astronomical Journal* **79** (1974) 745.
- [7] W.H. Richardson, *Bayesian-Based Iterative Method of Image Restoration*, *Journal of the Optical Society of America* (1917-1983) **62** (1972) 55.
- [8] C. Karwin, S. Boggs, J. Tomsick, A. Zoglauer, H. Lazar, J. Beechert et al., *The COSI Data Challenges and Simulations*, in *AAS/High Energy Astrophysics Division*, vol. 54 of *AAS/High Energy Astrophysics Division*, p. 108.30, Apr., 2022.